# Pauli channel tomography with unknown channel directions

Dániel Virosztek[2], László Ruppert[1,2], Katalin M. Hangos[1]

[1] Process Control Research Group,
Computer and Automation Research Institute,
H-1518 Budapest, POB 63, Hungary

[2] Department of Analysis,
Budapest University of Technology and Economics,
H-1521 Budapest, POB 91, Hungary

## Abstract

In this paper we estimate the parameters of the qubit Pauli channel using the channel matrix formalism. The main novelty of this work is that we do not assume the directions of the Pauli channel to be known, but they are determined through the tomography process, too. The results show that for optimally estimating the contraction parameters and the channel matrix we should have input qubits and measurements in the channel directions. However, for optimally estimating the channel directions, we should use different tomography conditions.

**Keywords:** experiment design, Pauli channel, parameter estimation, measurement, quadratic error, qubit.

# 1  Introduction

The accurate description of different quantum phenomena is a key issue in their potential use in modern IT-applications. In quantum mechanics, both dynamical changes and communication is treated using quantum channels. Therefore the parameter estimation of quantum channels plays a major role in quantum information processing, and the area of quantum process tomography is flourishing [1, 4, 8, 15, 18].

Direct quantum process tomography is performed by sending known quantum systems into the channel, and then estimating the output state. In quantum mechanics the measurement has a probabilistic nature [11, 12], therefore many identical copies of the input quantum system are needed, and an estimator is constructed by using statistical considerations. For achieving efficient process tomography, experiment design is necessary that

consists of selecting the optimal input state, optimal measurement of the output state, and an efficient estimator of the channel from the measured data.

The field of quantum process tomography is well-established, an exhaustive description of possible tomography methods can be found in [9]. The Pauli channels form a relatively wide family of quantum channels. The tomography of Pauli channels has a huge literature, however, due to the level of difficulty of the topic, papers mostly deal with special cases, e.g., with the optimal parameter estimation of a depolarizing channel [17]. But there are some publications investigating the estimation of multi-parameter channels [4, 19], and the multidimensional case also appears [7, 10]. There are also some experimental results concerning the optimal estimation of the Pauli-channels [5, 6].

In contrast to the majority of the works in this area, we propose an extended problem statement: we investigate qubit Pauli channels with unknown channel directions. Despite of the novelty of the approach, there are a few papers that deal with optimally estimating qubit Pauli channels including their channel directions. In [2] the problem was examined using convex optimization methods, and a numerical method was provided for finding the optimal input - measurement pairs. In [16] we examined the optimality of the estimation problem using purely statistical considerations to achieve analytical results. However, analytical results could only be obtained for the case of known channel directions.

Therefore, the aim of this paper is to give an analytical description of the optimal estimation of Pauli channels in the case of unknown channel directions, too. The efficiency of these estimations is measured here with three quantities: the mean squared error of the estimated contraction parameters and angle parameters, and the mean distance of the estimated and the real channel matrix are investigated.

The paper is organized as follows. In Section 2 we introduce the necessary notions to understand the rest of the article. In Section 3 the tomography method used is described. In Section 4 the optimization of the previously mentioned quantities are performed. Finally, conclusions are drawn.

## 2  Preliminaries

In the following section, we give a short introduction to the applied concepts. A more detailed description can be found in [11, 12]. We will only examine two-level systems, that is quantum bits or qubits.

The state of two-level quantum systems is described by *density matrices* $\rho \in \mathbf{M}_2(\mathbb{C})$ that are parametrized by a real vector $\theta \in \mathbb{R}^3$ called the *Bloch vector*.

**Definition 1** (Bloch parametrization)**.**

$$\rho : \mathbb{R}^3 \to \mathbf{M}_2(\mathbb{C}); \ \theta = (\theta_1, \theta_2, \theta_3)^T \mapsto \rho(\theta) = \frac{1}{2} \left( I + \theta_1 \sigma_1 + \theta_2 \sigma_2 + \theta_3 \sigma_3, \right) \qquad (2.1)$$

*where*

$$I = \sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \ \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \ \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

**Definition 2** (Quantum state)**.** *A qubit can be described with $2 \times 2$ density matrices $(\rho(\theta))$ satisfying the following condition:*

$$\text{Tr}(\rho(\theta)) = 1$$
$$\rho(\theta) \geq 0.$$

It is easy to check that $\rho(\theta)$ is a density matrix if and only if $\theta_1^2 + \theta_2^2 + \theta_3^2 \leq 1$. That is, the state space can be represented with the unit ball in $\mathbb{R}^3$, the so-called Bloch ball.

We will use *von Neumann measurements* with two possible outcomes [14].

**Definition 3** (Measurement). *$\{P, I - P\}$ is a von Neumann measurement if $P$ is $2 \times 2$ projection. The probability of measuring outcome $P$ on the system $\rho(\theta)$ is $\mathrm{Tr}(\rho(\theta)P)$.*

The only non-trivial case is when $P$ is a rank-one projection. Then using the same Bloch parametrization as previously, we can rewrite $P$ in the following form:

$$P = \frac{1}{2}\left(\sigma_0 + m_1\sigma_1 + m_2\sigma_2 + m_3\sigma_3\right), \tag{2.2}$$

where $m_1^2 + m_2^2 + m_3^2 = 1$. Using the abbreviation $\underline{m} = (m_1, m_2, m_3)^T$ and $\underline{\theta} = (\theta_1, \theta_2, \theta_3)^T$ we obtain

$$\mathrm{Prob}(\text{"measuring P"}) = \mathrm{Tr}\left(\rho(\theta)P\right) = \frac{1}{2}(1 + \underline{m} \cdot \underline{\theta}).$$

*Quantum channels* are completely positive, trace-preserving maps, and Pauli channels are a well-known family of them in the qubit case.

**Definition 4** (Pauli channel). *Let be $\{v_0 = I, v_1, v_2, v_3\}$ an arbitrary base satisfying $\mathrm{Tr}(v_i v_j) = 2\delta_{i,j}$ , $v_i \in \mathbf{M}_2^{s.a.}(\mathbb{C})$ $\forall i, j \in \{0, 1, 2, 3\}$. Let be $\lambda_1, \lambda_2, \lambda_3$ real numbers that fulfill*

$$1 \pm \lambda_3 \geq |\lambda_1 \pm \lambda_2|. \tag{2.3}$$

*Then a Pauli channel can be described with a mapping*

$$\mathcal{E} : \mathbf{M}_2(\mathbb{C}) \to \mathbf{M}_2(\mathbb{C}); \ \rho = \frac{1}{2}\left(I + \sum_{i=1}^{3}\theta_i v_i\right) \mapsto \mathcal{E}(\rho) = \frac{1}{2}\left(I + \sum_{i=1}^{3}\lambda_i\theta_i v_i\right). \tag{2.4}$$

*The affine subspaces $\{\frac{1}{2}(I + tv_i) : t \in \mathbb{R}\} \subset \mathbf{M}_2^{s.a.}(\mathbb{C})$ ($i \in \{1, 2, 3\}$) are called the channel directions, the numbers $\lambda_1, \lambda_2, \lambda_3$ are called the contraction parameters.*

The completely positiveness of the mapping is guaranteed by condition (2.3) (see [13]). From (2.3) follows that $-1 \leq \lambda_i \leq 1$, $i \in \{1, 2, 3\}$. If we look at the effect of a Pauli channel on the Bloch vectors then it becomes clear that we have contractions with parameter $\lambda_i$ in the appropriate channel directions. In other words, the image of the whole Bloch ball will be an ellipsoid with its axes lying in the directions of the channel and being $2|\lambda_i|$ long.

**Definition 5** (Channel matrix). *We will call the mapping $A : \mathbb{R}^3 \to \mathbb{R}^3$ the channel matrix of Pauli channel $\mathcal{E}$, if*

$$\mathcal{E} \circ \rho = \rho \circ A. \tag{2.5}$$

Straightforward calculations show that $A$ is a linear mapping and it can be always parametrized in the following form:

**Theorem 1.** *The channel matrix A of every Pauli channel can be constructed as*

$$A(\lambda_1, \lambda_2, \lambda_3, \phi_z, \phi_y, \phi_x) = R_z R_y R_x \Lambda R_x^{-1} R_y^{-1} R_z^{-1}, \tag{2.6}$$

*where*

$$R_z(\phi_z) = \begin{pmatrix} \cos\phi_z & -\sin\phi_z & 0 \\ \sin\phi_z & \cos\phi_z & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad R_y(\phi_y) = \begin{pmatrix} \cos\phi_y & 0 & -\sin\phi_y \\ 0 & 1 & 0 \\ \sin\phi_z & 0 & \cos\phi_y \end{pmatrix},$$

$$R_x(\phi_x) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi_x & -\sin\phi_x \\ 0 & \sin\phi_x & \cos\phi_x \end{pmatrix}, \qquad \Lambda = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}.$$

Note that this parametrization is only surjective, not bijective. This means that we can get the same channel matrix for several set of parameters.

# 3 Tomography of Pauli channels

Using the notions described in the previous section, we can characterize the Pauli channel with its channel matrix ($A$). Furthermore, Theorem 1 shows that the channel matrix can be constructed by using three contraction parameters ($\lambda_i$) and three angle parameters ($\phi_i$). Our aim is to give the best estimation of these quantities.

For this purpose we will send some input qubits through the channel, perform some measurements and construct an estimator from the measured data.

**Input qubits and measurements** We have to have at least three different input qubits for complete channel tomography. Let us suppose that we have three different measurements, too. Previous investigations showed that one should choose pure input states, which are orthogonal to each other [2],[16].

Let us suppose that the Bloch vectors of the input states are $\underline{\theta}^{(1)}, \underline{\theta}^{(2)}, \underline{\theta}^{(3)}$. If they are orthogonal, they create an orthogonal matrix and we can parametrize them the following way

$$\Theta = \left[\underline{\theta}^{(1)}, \underline{\theta}^{(2)}, \underline{\theta}^{(3)}\right] = R_z(\vartheta_z) R_y(\vartheta_y) R_x(\vartheta_x), \tag{3.1}$$

where $R_z, R_y, R_x$ are the same rotations as in (2.6) and $0 \leq \vartheta_z, \vartheta_y < \pi$, $0 \leq \vartheta_x < \frac{\pi}{2}$.

Similarly, we suppose that the measurements with Bloch vectors $\underline{m}^{(1)}, \underline{m}^{(2)}, \underline{m}^{(3)}$ are orthogonal [16]. Thus, we can parametrize them using one-dimensional rotations, too:

$$M = \left[\underline{m}^{(1)}, \underline{m}^{(2)}, \underline{m}^{(3)}\right] = R_z(\tau_z) R_y(\tau_y) R_x(\tau_x) \tag{3.2}$$

where $0 \leq \tau_z, \tau_y < \pi$, $0 \leq \tau_x < \frac{\pi}{2}$.

**Estimation of the channel matrix** The matrix generated from the Bloch vectors of the output qubits $\Xi = \left[\underline{\xi}^{(1)}, \underline{\xi}^{(2)}, \underline{\xi}^{(3)}\right]$ fulfill

$$\Xi = A(\lambda_1, \lambda_2, \lambda_3, \phi_z, \phi_y, \phi_x)\Theta. \tag{3.3}$$

Let us suppose that we perform the $i$-th measurement $N$ times on the copies of the $j$-th output qubit and denote by $N_{ij}$ the number of measurement results corresponding to $m^{(i)}$. Then $N_{ij}$ is binomially distributed with the following parameters

$$N_{ij} \sim \text{Binom}\left(N, \frac{1 + \underline{m}^{(i)} \cdot \underline{\xi}^{(j)}}{2}\right). \tag{3.4}$$

We can use the notation $x_{ij} = \underline{m}^{(i)} \cdot \underline{\xi}^{(j)}$, $\forall\, i, j \in \{1, 2, 3\}$, then $X = \{x_{ij}\}_{i,j=1}^{3}$ can be written in the form

$$X = M^{T}\Xi. \tag{3.5}$$

From (3.4) we can estimate the elements of $X$ (since $\mathbf{E}(N_{ij}) = N \cdot \frac{1 + x_{ij}}{2}$):

$$\hat{x}_{ij} := \frac{2}{N} \cdot N_{ij} - 1. \tag{3.6}$$

Furthermore, the elements of $\Xi$ can also be estimeted from (3.5) as

$$\hat{\Xi} := M\hat{X}. \tag{3.7}$$

Finally, from (3.3) we can obtain an estimation of the channel matrix

$$\hat{A} := \hat{\Xi}\Theta^{-1} = M\hat{X}\Theta^{T}. \tag{3.8}$$

**Estimation of the channel parameters**   From (2.6) we know the mapping from the sets of parameters to the channel matrices:

$$A : \mathcal{D} \to \mathbf{M}_3(\mathbb{R}); \ (\lambda_1, \lambda_2, \lambda_3, \phi_z, \phi_y, \phi_x) \mapsto A(\lambda_1, \lambda_2, \lambda_3, \phi_z, \phi_y, \phi_x),$$

where $\mathcal{D} \subset \mathbb{R}^6$, and our aim is to define its inverse mapping, i.e., we want to estimate the channel parameters from the estimation of the channel matrix (3.8). That is we need to find the mapping

$$T : \mathbf{M}_3(\mathbb{R}) \to \mathbb{R}^6; \ \hat{A} \mapsto (\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\phi}_z, \hat{\phi}_y, \hat{\phi}_x),$$

which fulfills

$$T \circ A = \text{Id}_{\mathcal{D}}. \tag{3.9}$$

We can construct $T$ the following way. Let us symmetrize our estimation:

$$\hat{A}_s := \frac{1}{2}\left(\hat{A} + \hat{A}^{T}\right).$$

Then $\hat{A}_s$ can be diagonalized in an orthonormal basis. The Jordan decomposition of $\hat{A}_s$ can be calculated easily algebraically since its characteristic polynomial is cubic. This way we get the eigenvalues $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \hat{\lambda}_3$, which are exactly the estimates of the channel contraction parameters, and the corresponding normalized eigenvectors: $\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}$ are also obtained.

The next question is how we can calculate the angle parameters from these eigenvectors. From (2.6) we can see the geometrical meaning of the angle parameters: $\phi_z$ and $\phi_y$ are the polar and azimuth angles of $\mathbf{v_1}$, while $\phi_x$ is the angle of $\mathbf{v_2}$ and the intersection of planes $z = 0$ and $\mathbf{v_1}^{\perp}$ (orthogonal subspace of $\mathbf{v_1}$).

These quantities can be uniquely determined only if $\lambda_1 > \lambda_2 > \lambda_3$, so we need to restrict the domain of mapping $A$ ($\mathcal{D}$) to get a one-to-one correspondence between channel matrices and channel parameters.

**Theorem 2.** *The parametrization will be bijective on the following domain:*

$$\mathcal{D} = \{(\lambda_1, \lambda_2, \lambda_3, \phi_z, \phi_y, \phi_x) \in \mathbb{R}^6 \ : \ 1 \pm \lambda_3 \geq |\lambda_1 \pm \lambda_2|, \ \lambda_1 \geq \lambda_2 \geq \lambda_3,$$

$$\phi_z, \phi_y, \phi_x \in [0, \pi), \phi_y = \frac{\pi}{2} \Rightarrow \phi_z = 0, \ \lambda_1 = \lambda_2 = \lambda_3 \Rightarrow \phi_z = \phi_y = \phi_x = 0,$$

$$\lambda_1 = \lambda_2 > \lambda_3 \Rightarrow (\phi_x = 0 \ and \ \phi_y = 0 \Rightarrow \phi_z = 0),$$

$$\lambda_1 > \lambda_2 = \lambda_3 \Rightarrow \phi_x = 0\}.$$

# 4 Optimal input states and measurements

Let us introduce some abbreviations:

$$\underline{\lambda} = [\lambda_1, \lambda_2, \lambda_3] \ - \ \text{channel contraction parameters},$$
$$\underline{\phi} = [\phi_z, \phi_y, \phi_x] \ - \ \text{channel angle parameters},$$
$$\underline{\tau} = [\tau_z, \tau_y, \tau_x] \ - \ \text{measurement parameters},$$
$$\underline{\vartheta} = [\vartheta_z, \vartheta_y, \vartheta_x] \ - \ \text{input qubits parameters}.$$

In the previous section we gave an estimation method of the channel matrix $(A)$, contraction parameters $(\underline{\lambda})$ and angle parameters $(\underline{\phi})$, now we want to analyze the estimates.

The distance between the real and estimated parameters can be given in various ways [3]. Let us define our loss functions as the averaged squared errors of the previously mentioned quantities:

$$f(\underline{\lambda}, \underline{\phi}, \underline{\tau}, \underline{\vartheta}, N) = \mathbf{E}\left(\|A(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\phi}_z, \hat{\phi}_y, \hat{\phi}_x) - A(\lambda_1, \lambda_2, \lambda_3, \phi_z, \phi_y, \phi_x)\|^2\right), \qquad (4.1)$$

$$g(\underline{\lambda}, \underline{\phi}, \underline{\tau}, \underline{\vartheta}, N) = \mathbf{E}\left((\hat{\lambda}_1 - \lambda_1)^2 + (\hat{\lambda}_2 - \lambda_2)^2 + (\hat{\lambda}_3 - \lambda_3)^2\right), \qquad (4.2)$$

$$h(\underline{\lambda}, \underline{\phi}, \underline{\tau}, \underline{\vartheta}, N) = \mathbf{E}\left(\text{dist}(\hat{\phi}_z, \phi_z)^2 + \text{dist}(\hat{\phi}_y, \phi_y)^2 + \text{dist}(\hat{\phi}_x, \phi_x)^2\right), \qquad (4.3)$$

where $\|\cdot\|$ is the Hilbert–Schmidt norm, $\text{dist}(\hat{\phi}, \phi) := \inf\{|\hat{\phi} - (\phi + k\pi)| : \ k \in \mathbb{Z}\}$.

**Remark.** *The estimation error of the channel matrix is equivalent to the average estimation error of output qubits, since*

$$\int_\theta \|\mathcal{E}(\rho(\theta)) - \hat{\mathcal{E}}(\rho(\theta))\|^2 d\theta = \int_\theta \frac{1}{2}\|A \ \theta - \hat{A} \ \theta\|^2 d\theta = \frac{1}{2}\int_\theta \sum_{i=1}^3 \left(\sum_{j=1}^3 (A_{ij} - \hat{A}_{ij})\theta_j\right)^2 d\theta =$$

$$= \frac{1}{2}\int_\theta \sum_{i=1}^3 \sum_{j=1}^3 \theta_j^2 (A_{ij} - \hat{A}_{ij})^2 d\theta + \frac{1}{2}\int_\theta \sum_{i=1}^3 \sum_{j<k} 2\theta_j \theta_k (A_{ij} - \hat{A}_{ij})(A_{ik} - \hat{A}_{ik}) d\theta =$$

$$= \frac{1}{2}\sum_{i=1}^3 \sum_{j=1}^3 \left(\int_\theta \theta_j^2 d\theta\right)(A_{ij} - \hat{A}_{ij})^2 + \sum_{i=1}^3 \sum_{j<k} \left(\int_\theta \theta_j \theta_k d\theta\right)(A_{ij} - \hat{A}_{ij})(A_{ik} - \hat{A}_{ik}) =$$

$$= c\|A - \hat{A}\|^2,$$

*where in the last step we used the symmetry of Bloch-ball: $\int_\theta \theta_j^2 d\theta =$ constant, $\int_\theta \theta_j \theta_k d\theta = 0$.*

In the following we will find the optimal inputs and measurements for a given channel $(\underline{\lambda}, \underline{\phi})$ and number of measurements $(N)$ that minimize the above loss functions.

## 4.1 The optimal estimation of the channel matrix

Before obtaining the optimal estimation settings, we will introduce some useful statements.

**Proposition 3** (Rotational invariance)**.** *For any $O \in \mathbf{M}_3(\mathbb{R})$ orthogonal matrix, the estimation of the Pauli channel described by the channel matrix $A$ using the input and measurement settings $\Theta$ and $M$ (see (3.1), (3.2)) is exactly as efficient as the estimation of the Pauli channel described by $OAO^{-1}$ with the input and measurement settings $O\Theta$ and $OM$.*

Therefore, it is enough to investigate Pauli channels with channel parameters $\phi_z = \phi_y = \phi_x = 0$.

**Lemma 4** (Unbiasedness 1)**.** *The estimation of the channel matrix given in (3.8) is unbiased.*

*Proof.* The estimation of the elements of $X$ defined in (3.6) is unbiased (that is, $\mathbf{E}\hat{X} = X$). The estimator of (3.8) is a linear function of $\hat{X}$, the expected value is linear, too, so $\mathbf{E}\hat{A} = A$. $\qquad\square$

Let us introduce the notation $\hat{a}_{ij} = [\hat{A}]_{ij}$ $(i,j \in \{1,2,3\})$ and the index set $H = \{11, 12, \ldots, 32, 33\}$. $\hat{A}$ is a linear transformation of $\hat{X}$, so

$$\hat{a}_k = \sum_{l \in H} c_{kl}(\underline{\tau}, \underline{\vartheta})\hat{x}_l \ (k \in H), \tag{4.4}$$

where the constants come from the actual values of $\Theta$ and $M$, hence determined by the parameters $\underline{\vartheta}$ and $\underline{\tau}$.

**Lemma 5.** *Set $\psi = \sum_{k \in H} d_k \hat{a}_k$ $(d_k \in \mathbb{R})$. Then*

$$\mathbf{Var}\,(\psi) = \sum_{l \in H} \left( \sum_{k \in H} d_k c_{kl}\,(\underline{\tau}, \underline{\vartheta}) \right)^2 \frac{1 - x_l^2}{N}. \tag{4.5}$$

*Proof.* From (3.4) and (3.6) it is easy to calculate the distribution of the elements of $\hat{X}$. Using the well known properties of the binomial distribution and the independence of the different elements of $\hat{X}$, straightforward calculations verify the statement. $\qquad\square$

**Proposition 6.** *The matrix $\{c_{kl}\}_{k,l \in H} \in \mathbf{M}_9(\mathbb{R})$ defined by equation (4.4) is orthogonal, because $\Theta$ and $M$ are orthogonal matrices. From this follows that the Hadamard-square of this matrix, $\{c_{kl}^2\}_{k,l \in H}$ is bistochastic. That is $\sum_{k \in H} c_{kl}^2 = 1 \ \forall l \in H$, $\sum_{l \in H} c_{kl}^2 = 1 \ \forall k \in H$.*

**Theorem 7.**
$$f(\underline{\lambda}, \underline{0}, \underline{\tau}, \underline{\vartheta}, N) \geq \frac{1}{N} \left( 6 - (\lambda_1^2 + \lambda_2^2 + \lambda_3^2) \right), \tag{4.6}$$

*and (4.6) holds with equality, if $\underline{\tau} = \underline{\vartheta} = \underline{0}$.*

*Proof.* The distance of the channel matrix and its estimation:

$$\|A(\lambda_1, \lambda_2, \lambda_3, \phi_z, \phi_y, \phi_x) - A(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\phi}_z, \hat{\phi}_y, \hat{\phi}_x)\|^2 = \|A(\lambda_1, \lambda_2, \lambda_3, \phi_z, \phi_y, \phi_x) - \hat{A}_s\|^2 =$$

$$= (\hat{a}_{11} - a_{11})^2 + (\hat{a}_{22} - a_{22})^2 + (\hat{a}_{33} - a_{33})^2 + 2\left(\frac{\hat{a}_{12} - a_{12}}{2} + \frac{\hat{a}_{21} - a_{21}}{2}\right)^2 +$$

$$+ 2\left(\frac{\hat{a}_{13} - a_{13}}{2} + \frac{\hat{a}_{31} - a_{31}}{2}\right)^2 + 2\left(\frac{\hat{a}_{23} - a_{23}}{2} + \frac{\hat{a}_{32} - a_{32}}{2}\right)^2.$$

From Lemma 4 the mean squared error of the $\hat{a}_{ij}$-s can be written as their variance, hence

$$f(\underline{\lambda}, \underline{0}, \underline{\tau}, \underline{\vartheta}, N) = \mathbf{E}\left(\|A(\lambda_1, \lambda_2, \lambda_3, \phi_z, \phi_y, \phi_x) - A(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\phi}_z, \hat{\phi}_y, \hat{\phi}_x)\|^2\right) =$$

$$= \frac{1}{2}\left(\mathbf{Var}(\hat{a}_{12} + \hat{a}_{21}) + \mathbf{Var}(\hat{a}_{13} + \hat{a}_{31}) + \mathbf{Var}(\hat{a}_{23} + \hat{a}_{32})\right) + \sum_{i \in \{1,2,3\}} \mathbf{Var}(\hat{a}_{ii}) =$$

$$= -\frac{1}{2}\left(\mathbf{Var}(\hat{a}_{12} - \hat{a}_{21}) + \mathbf{Var}(\hat{a}_{13} - \hat{a}_{31}) + \mathbf{Var}(\hat{a}_{23} - \hat{a}_{32})\right) + \sum_{i,j \in \{1,2,3\}} \mathbf{Var}(\hat{a}_{ij}).$$

By Lemma 5

$$\mathbf{Var}(\hat{a}_k) = \sum_{l \in H} c_{kl}^2(\underline{\tau}, \underline{\vartheta})\frac{1 - x_l^2}{N},$$

hence using the bistochastic property of the matrix $\{c_{kl}^2\}_{k,l \in H}$ described in Proposition 6, we can see that

$$\sum_{k \in H} \mathbf{Var}(\hat{a}_k) = \sum_{k \in H} \sum_{l \in H} c_{kl}^2(\underline{\tau}, \underline{\vartheta})\frac{1 - x_l^2}{N} = \sum_{l \in H} \frac{1 - x_l^2}{N} = \frac{1}{N}\left(9 - \sum_{l \in H} x_l^2\right). \quad (4.7)$$

On the other hand, Lemma 5 shows that

$$\mathbf{Var}\left(\hat{a}_m - \hat{a}_n\right) = \sum_{l \in H} (c_{ml} - c_{nl})^2 \frac{1 - x_l^2}{N}. \quad (4.8)$$

From the orthogonality of $\{c_{kl}\}_{k,l \in H}$ follows that $\sum_{l \in H}(c_{ml} - c_{nl})^2 = 2$, since this expression is the norm-square of orthogonal unit vectors in $\mathbb{R}^9$. Therefore $\frac{1 - x_l^2}{N} \leq \frac{1}{N}$ implies that $\mathbf{Var}\left(\hat{a}_m - \hat{a}_n\right) \leq \frac{2}{N}$.

Finally

$$\sum_{l \in H} x_l^2 = \mathrm{Tr}XX^T = \mathrm{Tr}(M^{-1}A\Theta)(\Theta^{-1}A^T M) = \mathrm{Tr}AA^T = \lambda_1^2 + \lambda_2^2 + \lambda_3^2 \quad (4.9)$$

hence

$$f(\underline{\lambda}, \underline{0}, \underline{\tau}, \underline{\vartheta}, N) \geq -\frac{1}{2}\left(\frac{2}{N} + \frac{2}{N} + \frac{2}{N}\right) + \frac{1}{N}\left(9 - (\lambda_1^2 + \lambda_2^2 + \lambda_3^2)\right).$$

which is equivalent to the inequality (4.6). It is easy to check that in the $\underline{\tau} = \underline{\vartheta} = \underline{0}$ case

$$\mathbf{Var}(\hat{a}_{ij}) = \frac{1}{N}(1 - \delta_{ij}\lambda_i^2), \ i, j \in \{1, 2, 3\},$$

$$\mathbf{Var}(\hat{a}_{12} - \hat{a}_{21}) = \mathbf{Var}(\hat{a}_{13} - \hat{a}_{31}) = \mathbf{Var}(\hat{a}_{23} - \hat{a}_{32}) = \frac{2}{N},$$

hence the minimum of $f(\underline{\lambda}, \underline{0}, \underline{\tau}, \underline{\vartheta}, N)$ is taken in $\underline{\tau} = \underline{\vartheta} = \underline{0}$. $\square$

## 4.2 The optimal estimation of the contraction parameters

It is difficult to compute the loss functions $g$ and $h$ defined in (4.2) and (4.3), because the $\lambda_i$ ($i \in \{1, 2, 3\}$) and the $\phi_\alpha$ ($\alpha \in \{z, y, x\}$) parameter estimators are non-linear. However, one can approximate these parameter estimators with the their first-order Taylor polynomial (recall that the $\hat{\lambda}_i, \hat{\phi}_\alpha$ estimators are $\mathbf{M}_3(\mathbb{R}) \to \mathbb{R}$ functions). Lemma 4 shows that $\mathbf{E}(\hat{a}_{ij}) = a_{ij}$, while from Lemma 5 it follows that $\mathbf{Var}(\hat{a}_{ij}) = \underline{O}\left(\frac{1}{N}\right)$. Hence the Taylor polynomial of functions $\hat{\lambda}_i, \hat{\phi}_\alpha$ with the base point $\mathbf{E}(\hat{A}) = A$ will be an appropriate approximation, if $N$ is large enough.

So the linearized estimators can be written in the form

$$\tilde{\lambda}_i := \hat{\lambda}_i(A) + \left\langle \operatorname{grad}\hat{\lambda}_i(A), \, \hat{A} - A \right\rangle \ (\forall \, i \in \{1, 2, 3\}), \tag{4.10}$$

$$\tilde{\phi}_\alpha := \hat{\phi}_\alpha(A) + \left\langle \operatorname{grad}\hat{\phi}_\alpha(A), \, \hat{A} - A \right\rangle \ (\forall \, \alpha \in \{z, y, x\}). \tag{4.11}$$

**Lemma 8** (Unbiasedness 2)**.** *The above defined $\tilde{\lambda}_i$ ($i \in \{1, 2, 3\}$), $\tilde{\phi}_\alpha$ ($\alpha \in \{z, y, x\}$) estimations of the channel parameters are unbiased.*

*Proof.* It follows from (3.9) that we have $\hat{\lambda}_i\left(A(\underline{\lambda}, \underline{\phi})\right) = \lambda_i$ and $\hat{\phi}_\alpha\left(A(\underline{\lambda}, \underline{\phi})\right) = \phi_\alpha$. Using the unbiasedness of $\hat{A}$ we get

$$\mathbf{E}(\tilde{\lambda}_i) = \mathbf{E}(\hat{\lambda}_i(A)) + \mathbf{E}\left(\left\langle \operatorname{grad}\hat{\lambda}_i(A), \, \hat{A} - A \right\rangle\right) = \hat{\lambda}_i(A) + \left\langle \operatorname{grad}\hat{\lambda}_i(A), \, \mathbf{E}(\hat{A} - A) \right\rangle =$$

$$= \hat{\lambda}_i(A) + 0 = \lambda_i,$$

and similarly

$$\mathbf{E}(\tilde{\phi}_\alpha) = \mathbf{E}(\hat{\phi}_\alpha(A)) + \mathbf{E}\left(\left\langle \operatorname{grad}\hat{\phi}_\alpha(A), \, \hat{A} - A \right\rangle\right) = \phi_\alpha.$$

$\square$

According to Proposition 3, it is enough to consider the case when the Pauli channel described by $A$ has angle parameters $\phi_z = \phi_y = \phi_x = 0$.

Let us define the mapping T with formula

$$T : \mathbf{M}_3(\mathbb{R}) \to \mathbb{R}^6; \ \hat{A} \mapsto (\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\phi}_z, \hat{\phi}_y, \hat{\phi}_x)$$

**Lemma 9.** *The non-vanishing components of* $\mathrm{d}T\left(A(\underline{\lambda}, \underline{0})\right)$ *are*

$$\frac{\partial \hat{\lambda}_1}{\partial \hat{a}_{11}} = \frac{\partial \hat{\lambda}_2}{\partial \hat{a}_{22}} = \frac{\partial \hat{\lambda}_3}{\partial \hat{a}_{33}} = 1, \ \frac{\partial \hat{\phi}_z}{\partial \hat{a}_{12,s}} = \frac{1}{\lambda_1 - \lambda_2}, \ \frac{\partial \hat{\phi}_y}{\partial \hat{a}_{13,s}} = \frac{1}{\lambda_1 - \lambda_3}, \ \frac{\partial \hat{\phi}_x}{\partial \hat{a}_{23,s}} = \frac{1}{\lambda_2 - \lambda_3},$$

*where $\hat{a}_{ij,s} = \frac{1}{2}\left(\hat{a}_{ij} + \hat{a}_{ji}\right)$.*

*Proof.* The parameter estimation $T$ is the left-inverse of the channel parametrization

$$A : \mathcal{D} \to \mathbf{M}_3(\mathbb{R}); \ (\lambda_1, \lambda_2, \lambda_3, \phi_z, \phi_y, \phi_x) \mapsto A(\lambda_1, \lambda_2, \lambda_3, \phi_z, \phi_y, \phi_x),$$

hence

$$\mathrm{d}T\left(A(\underline{\lambda}, \underline{\phi})\right) = \left(\mathrm{d}A(\underline{\lambda}, \underline{\phi})\right)^{-1}. \tag{4.12}$$

We can calculate the elements of $dA(\underline{\lambda}, \underline{\phi})$ easily. For example:

$$A(\lambda_1, \lambda_2, \lambda_3, \phi_z, 0, 0) = \begin{pmatrix} \lambda_1 \cos^2 \phi_z + \lambda_2 \sin^2 \phi_z & (\lambda_1 - \lambda_2) \sin \phi_z \cos \phi_z & 0 \\ (\lambda_1 - \lambda_2) \sin \phi_z \cos \phi_z & \lambda_1 \sin^2 \phi_z + \lambda_2 \cos^2 \phi_z & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

so

$$\frac{\partial A}{\partial \phi_z}(\lambda_1, \lambda_2, \lambda_3, 0, 0, 0) = \begin{pmatrix} 0 & (\lambda_1 - \lambda_2) & 0 \\ (\lambda_1 - \lambda_2) & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

From similar calculations we get that

$$dA(\lambda_1, \lambda_2, \lambda_3, 0, 0, 0) = \mathrm{Diag}(1, 1, 1, \lambda_1 - \lambda_2, \lambda_1 - \lambda_3, \lambda_2 - \lambda_3).$$

from which the statement follows. $\qquad\square$

Let us substitute the result of Lemma 9 into the definition of $\tilde{\lambda}_i$ (see (4.10)), and use the unbiasedness of $\tilde{\lambda}_i$, and $\hat{A}$. Then the loss function defined by the *linearized* parameter estimators will have the following form:

$$\tilde{g}(\underline{\lambda}, \underline{0}, \underline{\tau}, \underline{\vartheta}, N) := \mathbf{E}\left((\tilde{\lambda}_1 - \lambda_1)^2 + (\tilde{\lambda}_2 - \lambda_2)^2 + (\tilde{\lambda}_3 - \lambda_3)^2\right) =$$
$$= \mathbf{Var}(\hat{a}_{11}) + \mathbf{Var}(\hat{a}_{22}) + \mathbf{Var}(\hat{a}_{33}). \qquad (4.13)$$

This function is asymptotically equal to $g$, but easier to handle, thus, we can perform the analytical optimization.

**Theorem 10.**
$$\tilde{g}(\underline{\lambda}, \underline{0}, \underline{\tau}, \underline{\vartheta}, N) \geq \frac{1}{N}\left(3 - (\lambda_1^2 + \lambda_2^2 + \lambda_3^2)\right), \qquad (4.14)$$

and (4.14) holds with equality, if $\underline{\tau} = \underline{\vartheta} = 0$.

*Proof.*

$$\tilde{g}(\underline{\lambda}, \underline{0}, \underline{\tau}, \underline{\vartheta}, N) = \sum_{i \in \{1,2,3\}} \mathbf{Var}(\hat{a}_{ii}) = \sum_{i,j \in \{1,2,3\}} \mathbf{Var}(\hat{a}_{ij}) - \sum_{i \neq j} \mathbf{Var}(\hat{a}_{ij}). \qquad (4.15)$$

(4.5) shows that $\mathbf{Var}(\hat{a}_{ij}) \leq \frac{1}{N}$ ($\forall i, j$), since $\frac{1 - x_{ij}^2}{N} \leq \frac{1}{N}$ and the $\{c_{kl}^2\}_{k,l \in H}$ matrix is bistochastic. Therefore

$$\sum_{i \neq j} \mathbf{Var}(\hat{a}_{ij}) \leq \frac{6}{N}.$$

Now, by the results of (4.7) and (4.9) we can write the following inequality based on the equation (4.15):

$$\mathbf{Var}(\hat{a}_{11}) + \mathbf{Var}(\hat{a}_{22}) + \mathbf{Var}(\hat{a}_{33}) \geq \frac{1}{N}\left(9 - (\lambda_1^2 + \lambda_2^2 + \lambda_3^2)\right) - \frac{6}{N},$$

and this is the statement of the theorem. It is easy to see that in the $\underline{\tau} = \underline{\vartheta} = \underline{0}$ case $\mathbf{Var}(\hat{a}_{11}) = \frac{1}{N}(1 - \lambda_1^2)$, $\mathbf{Var}(\hat{a}_{22}) = \frac{1}{N}(1 - \lambda_2^2)$, $\mathbf{Var}(\hat{a}_{33}) = \frac{1}{N}(1 - \lambda_3^2)$, hence $\tilde{g}(\underline{\lambda}, \underline{0}, \underline{\tau}, \underline{\vartheta}, N)$ is minimal in $\underline{\tau} = \underline{\vartheta} = 0$. $\qquad\square$

## 4.3 The optimal estimation of the angle parameters

From Lemma 9 we can determine the loss function that measures the accuracy of the estimation of the angle parameters:

$$\tilde{h}(\underline{\lambda}, \underline{0}, \underline{\tau}, \underline{\vartheta}, N) = \mathbf{E}\left((\tilde{\phi}_z - \phi_z)^2 + (\tilde{\phi}_y - \phi_y)^2 + (\tilde{\phi}_x - \phi_x)^2\right) =$$

$$= \frac{1}{4(\lambda_1 - \lambda_2)^2}\mathbf{Var}(\hat{a}_{12} + \hat{a}_{21}) + \frac{1}{4(\lambda_1 - \lambda_3)^2}\mathbf{Var}(\hat{a}_{13} + \hat{a}_{31}) + \frac{1}{4(\lambda_2 - \lambda_3)^2}\mathbf{Var}(\hat{a}_{23} + \hat{a}_{32}).$$

Using the result of Lemma 5, $\mathbf{Var}(\hat{a}_{12} + \hat{a}_{21})$, $\mathbf{Var}(\hat{a}_{13} + \hat{a}_{31})$ and $\mathbf{Var}(\hat{a}_{23} + \hat{a}_{32})$ can be expressed as a function of $\underline{\tau}, \underline{\vartheta}$ and $X = \{x_l\}_{l \in H}$. The definition of matrices $\Theta$ and $M$ and equations (3.3) and (3.5) show that $X$ can be written as a function of $\underline{\tau}, \underline{\vartheta}$ and $\underline{\lambda}$ (recall that we fixed $\underline{\phi} = \underline{0}$). Hence, $\tilde{h}(\underline{\lambda}, \underline{0}, \underline{\tau}, \underline{\vartheta}, N)$ can be written in a quite extensive, but explicit, closed form.

For fixed $\underline{\lambda}$ and $N$, the optimization of $\tilde{h}(\underline{\lambda}, \underline{0}, \underline{\tau}, \underline{\vartheta}, N)$ is a minimization problem with six variables $(\tau_z, \tau_y, \tau_x, \vartheta_z, \vartheta_y, \vartheta_x)$, but unfortunately this optimization problem can not be solved analytically.

However, we can formulate conjectures based on numerical optimization computations.

**Conjecture 11.** *For any fixed parameters $\underline{\lambda}, N$*

- *if $\tilde{h}(\underline{\lambda}, \underline{0}, \underline{\tau}, \underline{\vartheta}, N)$ is minimal at $(\underline{\tau}_{opt}, \underline{\vartheta}_{opt})$, then $\underline{\tau}_{opt} = \underline{\vartheta}_{opt}$,*

- *the estimation strategies described by parameters $\underline{\tau}_1 = \underline{\vartheta}_1 = (\frac{\pi}{4}, \frac{\pi}{4}, 0)$ and $\underline{\tau}_2 = \underline{\vartheta}_2 = (\frac{\pi}{4}, 0, \frac{\pi}{4})$ are nearly optimal.*

The following examples taken from our numerical optimizations studies indicate the validity of Conjecture 11.

**Example 12.** *Let us fix the following parameters: $\lambda_1 = 0.8, \lambda_2 = 0.65, \lambda_3 = 0.5$ and $N = 1000$. Then the optimal $(\underline{\tau}_{opt}, \underline{\vartheta}_{opt})$ (we can calculate them numerically) does not show any regularity except $\underline{\tau}_{opt} = \underline{\vartheta}_{opt}$, with*

$$\min_{\underline{\tau}, \underline{\vartheta}} \tilde{h} = \tilde{h}(\underline{\tau}_{opt}, \underline{\vartheta}_{opt}) = 0.03634.$$

*We can calculate the values at the two points given in Conjecture 11: $\tilde{h}(\underline{\tau}_1, \underline{\vartheta}_1) = \tilde{h}(\underline{\tau}_2, \underline{\vartheta}_2) = 0.03676$. So the difference can be considered small compared to $\tilde{h}(\underline{0}, \underline{0}) = 0.05$.*

**Example 13.** *The situation is similar for the parameters $\lambda_1 = 0.9, \lambda_2 = 0.67, \lambda_3 = 0.6$ and $N = 1000$. For the optimal input and measurement directions we have $\underline{\tau}_{opt} = \underline{\vartheta}_{opt}$ from numerical optimization, with*

$$\min_{\underline{\tau}, \underline{\vartheta}} \tilde{h} = \tilde{h}(\underline{\tau}_{opt}, \underline{\vartheta}_{opt}) = 0.01659.$$

*In comparison, $\tilde{h}(\underline{\tau}_1, \underline{\vartheta}_1) = \tilde{h}(\underline{\tau}_2, \underline{\vartheta}_2) = 0.01675$ and $\tilde{h}(\underline{0}, \underline{0}) = 0.02446$.*

**Pauli channel with known parameters in one direction** Let us assume that we have some information about the Pauli channel:

$$v_3 = \sigma_3 \text{ and } \lambda_3 = 0. \tag{4.16}$$

In this case the channel matrix has the following simplified form

$$A(\lambda_1, \lambda_2, \phi) = R(\phi)\Lambda(\lambda_1, \lambda_2)R(\phi)^{-1}. \tag{4.17}$$

Now its is easy to show that the input states and the von Neumann measurements should be orthogonal vectors in the plane spanned by $\sigma_1$ and $\sigma_2$. Hence, the input states and the measurements can be parametrized using a single angle parameter $\vartheta$ and $\tau$, respectively:

$$\Theta = \begin{bmatrix} \theta_{11} & \theta_{21} & 0 \\ \theta_{12} & \theta_{22} & 0 \\ 0 & 0 & 1 \end{bmatrix} = R(\vartheta), \quad M = \begin{bmatrix} m_{11} & m_{21} & 0 \\ m_{12} & m_{22} & 0 \\ 0 & 0 & 1 \end{bmatrix} = R(\tau).$$

Then we have to solve the following minimization problem with two variables $(\tau, \vartheta)$ for fixed $\lambda_1, \lambda_2, N$ values:

$$\tilde{h}_2(\lambda_1, \lambda_2, 0, \tau, \vartheta, N) = \mathbf{E}\left(\tilde{\phi} - \phi\right)^2. \tag{4.18}$$

It can be solved analytically, but the proof is rather technical than difficult. The result is formulated in the following theorem:

**Theorem 14.** *Assume that $\lambda_1 > \lambda_2$ and $\lambda_1 \neq -\lambda_2$.*

1. *If $(\lambda_1 + \lambda_2)^2 \geq 2(\lambda_1 - \lambda_2)^2$, then $\tilde{h}_2(\lambda_1, \lambda_2, 0, \tau, \vartheta, N)$ is minimal if and only if*

$$\tau_{opt} = \vartheta_{opt} = \frac{\pi}{4} \quad \left(\text{mod } \frac{\pi}{2}\right).$$

   *The minimal value of the loss function is*

$$\tilde{h}_2\left(\lambda_1, \lambda_2, 0, \frac{\pi}{4}, \frac{\pi}{4}, N\right) = \frac{1}{4(\lambda_1 - \lambda_2)^2}\frac{1}{2N}\left(4 - (\lambda_1 + \lambda_2)^2\right).$$

2. *If $(\lambda_1 + \lambda_2)^2 < 2(\lambda_1 - \lambda_2)^2$, then $\tilde{h}_2(\lambda_1, \lambda_2, 0, \tau, \vartheta, N)$ is minimal if and only if*

$$\tau_{opt} = \vartheta_{opt} = x \text{ or } \frac{\pi}{2} - x \quad \left(\text{mod } \frac{\pi}{2}\right),$$

   *where $x = \frac{1}{4}\arccos\left(-\frac{(\lambda_1 + \lambda_2)^2}{2(\lambda_1 - \lambda_2)^2}\right)$. The minimal value is now*

$$\tilde{h}_2(\lambda_1, \lambda_2, 0, \tau_{opt}, \vartheta_{opt}, N) = \frac{1}{4(\lambda_1 - \lambda_2)^2}\frac{1}{2N}\left(4 - (\lambda_1^2 + \lambda_2^2) - \frac{1}{8}\frac{(\lambda_1 + \lambda_2)^4}{(\lambda_1 - \lambda_2)^2}\right).$$

**Example 15.** *If $\lambda_1 = 0.8$ and $\lambda_2 = 0.2$ the optimal input and measurement directions are $\tau_{opt} = \vartheta_{opt} = \frac{\pi}{4}$. These are the optimal angles in most cases, however, if $\lambda_1 = 1$ and $\lambda_2 = 0$ then $\tau_{opt} = \vartheta_{opt} = \frac{\pi}{3}$ or $\frac{\pi}{6}$.*

# 5 Conclusion and discussion

In Theorem 1 we gave a parametrization of the channel matrix $A$, while in Theorem 2 we obtained the domain where the parametrization is bijective. We gave an estimation procedure for the channel matrix and channel parameters in Section 3, and then optimized this process with respect to input qubit and measurement directions. We have proven that for the optimal estimation of a channel matrix, we have to take input qubit – measurement pairs in the channel directions (Theorem 7), and the same statement is true for estimating the channel contraction parameters (Theorem 10). This result can be implemented using a two-step algorithm for optimally estimating the channel parameters, where in the first, shorter step we make a rough estimation on the channel directions, and then we set these directions for the second step of the algorithm when the contraction parameters are obtained.

The estimation of the angle parameters can not be performed analytically in the general case, and the loss function has a quite complex form even in a simplified case. From simulation investigations we conjecture that using a complementary basis to the channel directions would give a nearly optimal result.

# References

[1] A. Anis and A. I. Lvovsky: Maximum-likelihood coherent-state quantum process tomography. *New J. Phys.* **14**, 105021, 2012.

[2] G. Balló, K. M. Hangos and D. Petz: Convex Optimization-Based Parameter Estimation and Experiment Design for Pauli Channels. *IEEE Trans. on Automatic Control* **57**(8), 2056-2061, 2012.

[3] V. Belavkin: Contravariant densities, complete distances and relative fidelities for quantum channels. *Rep. Math. Phys.* **55**(1), 61-77, 2005.

[4] A. Bendersky, F. Patawski, and J.P. Paz: Selective and efficient estimation of parameters for quantum process tomography. *Phys. Rev. Lett.* **100**, 190403, 2008.

[5] M.P.A. Branderhorst, J. Nunn, I.A. Walmsley, and R.L. Kosut: Simplified quantum process tomography. *New J. Phys.* **11**, 115010, 2009.

[6] A. Chiuri, V. Rosati, G. Vallone, S. Padua, H. Imai, S. Giacomini, C. Macchiavello, and P. Mataloni. Experimental realization of optimal noise estimation for a general Pauli channel. *Phys. Rev. Lett.*, **107**, 253602, 2011.

[7] A. Fujiwara and H. Imai: Quantum parameter estimation of a generalized Pauli channel. *J. Phys. A*, **36**, 8093–8103, 2003.

[8] S. Rahimi-Keshari, A. Scherer, A. Mann, A. T. Rezakhani, A. I. Lvovsky and B. C. Sanders: Quantum process tomography with coherent states. *New J. Phys.*, **13**, 013006 , 2011.

[9] M. Mohseni, A. T. Rezakhani, and D. A. Lidar: Quantum process tomography: Resource analysis of different strategies. *Phys. Rev. A*, **77**, 032322, 2008.

[10] M. Nathanson and M. B. Ruskai: Pauli diagonal channels constant on axes. *J. Phys. A*, **40**, 8171-8204, 2007.

[11] M. A. Nielsen and I. L. Chuang: *Quantum Computation and Quantum Information.* Cambridge University Press, 2000.

[12] D. Petz: *Quantum Information Theory and Quantum Statistics.* Theoretical and Mathematical Physics. Springer-Verlag, 2008.

[13] D. Petz and H. Ohno: Generalizations of Pauli channels. *Acta Math. Hungar.*, **124**, 165-177, 2009.

[14] D. Petz and L. Ruppert: Efficient quantum tomography needs complementary and symmetric measurements. *Rep. Math. Phys.*, **69**(2), 161-177, 2012.

[15] W. Roga, M. Fannes and K. Zyczkowski: Davies maps for qubits and qutrits. *Rep. Math. Phys.* **66**(3), 311-329, 2010.

[16] L. Ruppert, D. Virosztek and K. M. Hangos: Optimal parameter estimation of Pauli channels. *J. Phys. A*, **45**, 265305, 2012.

[17] M. Sasaki, M. Ban, and S. M. Barnett: Optimal parameter estimation of a depolarizing channel. *Phys. Rev. A*, **66**(2), 022308, 2002.

[18] Y. S. Teo, B-G. Englert, J. Rehacek and Z. Hradil: Adaptive schemes for incomplete quantum process tomography *Phys. Rev. A*, **84**, 062125 , 2011.

[19] K. C. Young, M. Sarovar, R. Kosut, and K. B. Whaley: Optimal quantum multi-parameter estimation and application to dipole- and exchange-coupled qubits. *Phys. Rev. A*, **79**(6), 062301, 2009.